

Serial No.: 10/826,204

REMARKS

Claims 1-19 are pending in the application. Claims 1-19 are allowable and prosecution on the merits is closes. This amendment is for purposes of correcting the following informalities:

The specification is objected to because of informalities. The informalities cited by the examiner have been corrected in the substitute specification. The abstract has been amended to be within 50 to 150 words.

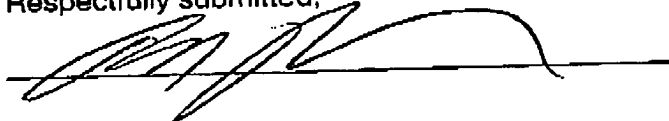
The claims are objected to due to mis-numbering of the claims. As indicated by the examiner, mis-numbered claim 16 has been renumbered to claim 16 such that the claims are numbered consecutively. No other amendments to the claims have been made

II. CONCLUSION

Accordingly, the application is believed to be allowable and the application is believed to be in condition for allowance. A prompt action to such end is earnestly solicited.

Should the Examiner feel that a telephone interview would be helpful to facilitate favorable prosecution of the above-identified application, the Examiner is invited to contact the undersigned at the telephone number provided below.

Respectfully submitted,

A handwritten signature in black ink, consisting of stylized, overlapping loops and strokes, positioned above a horizontal line.

Serial No.: 10/826,204

Timothy P. O'Hagan

Reg. No. 39,319

DATE:

1-4-08

Timothy P. O'Hagan
8710 Kilkenny Ct
Fort Myers, FL 33912
(239) 561-2300

10/826,204

Marked to Show Changes

Substitute Specification – Marked to Show Changes

TITLE: IMPROVED ADAPTIVE JITTER BUFFER SYSTEM AND JITTER CORRECTION METHOD FOR PACKET VOICE COMMUNICATIONS.

Technical Field

The present invention relates to packet voice communications, and more particularly to systems and methods which correct for variable latency in receipt of data packets containing compressed audio data.

Background of the Invention

For many years voice telephone service was implemented over a circuit switched network commonly known as the public switched telephone network (PSTN) and controlled by a local telephone service provider. In such systems, the analog electrical signals representing the conversation are transmitted between the two telephone handsets on a dedicated twisted-pair-copper-wire circuit. More specifically, each of the two endpoint telephones is coupled to a local switching station by a dedicated pair of copper wires known as a subscriber loop. The two switching stations are connected by a trunk line network comprising multiple copper wire pairs. When a telephone call is placed, the circuit is completed by dynamically coupling each subscriber loop to a dedicated pair of copper wires in the trunk line network.

Because each call is placed over a dedicated circuit, the delay in transmission of the audio signal is only the transmission latency of the dedicated circuit - which is typically imperceptible and remains relatively constant for the entire duration of the telephone call. Due to speech or other audio data being continuous in nature, an imperceptible and constant transmission delay is required to accurately reproduce the speech or other audio data at a receiving system.

10/826,204

Marked to Show Changes

29 More recently, the analog circuits between switching stations have been replaced
30 with digital transmission mediums which carry compressed digital audio data for multiple
31 telephone calls simultaneously. More specifically, at a first switching station the audio
32 may be digitized, compressed, and framed for transmission across the digital
33 transmission medium. At the receiving switching station, the frames are collected and
34 audio is reproduced. To avoid irregularity in the time of arrival of transmitted frames
35 (e.g. jitter) and gaps in the reproduced audio, a dedicated periodic time slot on the
36 transmission medium is reserved for each telephone call. In effect, the dedicated time
37 slot solution is equivalent to a dedicated circuit between the two stations.

38 More recently, Advances in the speed of data transmissions and Internet
39 bandwidth have made it possible for telephone conversations to be communicated
40 using the Internet's packet switched architecture with the overhead of Voice over
41 Internet Protocols (VoIP) such as the Real Time Protocol (RTP) and the UDP/IP
42 protocols.

43 In general VoIP utilizes network bandwidth more efficiently in that bandwidth on
44 any transmission segment may be utilized without reservation of dedicated time slots for
45 audio channels. Further, the routers of the Internet may route each frame from its
46 source to its destination based on real time segment usage.

47 A problem with use of VoIP for maintaining a telephone call between two stations
48 is that the transmission latency is not constant. The transmission time between when a
49 frame is released from the first station and received at the destination varies with each
50 frame. This variation is referred to as frame jitter. Further, frames may arrive out of
51 sequence or may not even arrive at all if the frame is lost in a buffer overflow at a router
52 along the Internet. This jitter and frame loss can cause gaps and clipping in the
53 reproduced audio.

54 To compensate for frame jitter, jitter buffers have been developed. In general, a
55 jitter buffer receives each frame from the transmission medium and then provides the
56 frames to a decompression circuit. While the frames may be received with variable

10/826,204

Marked to Show Changes

latency, the frames may be output to the decompression circuit at periodic intervals – so long as the jitter buffer does not empty or overflow. While a large jitter buffer with significant delay reduces the probability of the buffer becoming empty or overflowing, the significant delay itself degrades the quality of the telephone call.

To improve call quality, adaptive jitter buffers have been developed. In general, an adaptive jitter buffer increases the delay (and therefore the number of frames in the buffer) when jitter increases (increasing variation in frame latency) to assure that the buffer does not empty and decreases delay when jitter decreases (decreasing variation in frame latency) to decrease the overall delay between when the audio is spoken at the source station and reproduced at the receiving station.

Known adaptive jitter buffer systems are slow to react to changes in frame jitter. What is needed is an improved adaptive jitter buffer system and jitter correction method that does not suffer the reaction delays and other disadvantages of known systems.

Summary of the Invention

A first aspect of the present invention is an improved adaptive jitter buffer system for reducing jitter in a packet audio reception device such as a Voice over Internet Protocol (VoIP) telephone or terminal adapter.

The jitter buffer system comprises a jitter buffer, a delay calculation module, an output time stamp index module, and a histogram module – which, in the aggregate control a jitter buffer.

The jitter buffer stores a plurality of audio frames and provides for each of the plurality of audio frames to be released to a decompression circuit upon receipt of a signal therefore.

The delay calculation module receives each of the plurality of audio frames and, for each of the plurality of audio frames; i) calculates a delay value, ii) drops the frame if the delay value is less than zero; iii) drops the frame if the delay value is greater than a predetermined maximum delay value; and iv) writes the frame to the jitter buffer if the

10/826,204

Marked to Show Changes

85 delay value is greater than zero and less than the predetermined maximum delay value.
86 The delay value is equal to the time difference between the output time stamp value
87 and a transmission time stamp assigned to the frame by a transmitting system.

88 The output time stamp index module determines an initial output time stamp
89 following receipt of a jitter buffer latency value from the histogram module. The initial
90 output time stamp value is equal to the sum of a transmission time stamp assigned to a
91 first frame and the jitter buffer latency value. Thereafter, the output time stamp index
92 module increments the output time stamp value by a time period upon each release of a
93 from the jitter buffer to the decompression circuit.

94 The histogram module is coupled to each of the output time stamp index and the
95 delay calculation module. The histogram module periodically: i) calculates a target
96 delay value which, based on a buffered history of histogram values, would have resulted
97 in a predetermined portion of a fixed quantity of the most recently received frames being
98 dropped; ii) adjusting the jitter buffer latency value to a value equal to the target value;
99 and iii) providing the jitter buffer latency value to the output time stamp index module.

100 Each histogram value represents the delay value of each of the fixed quantity of
101 the most recently received frames. More specifically, the histogram value may be the
102 value of delay less the ~~then~~ current jitter buffer latency value.

103 In the exemplary embodiment, there exist rules regarding the adjustment of the
104 jitter buffer latency value. For example, the histogram module: i) only adjusts the jitter
105 buffer latency value to the target value if the difference between the jitter buffer latency
106 value and the target value is greater than a preconfigured hysteresis threshold; ii)
107 adjusts the jitter buffer latency value to a maximum preconfigured jitter buffer latency
108 value if the target delay value is greater than the maximum preconfigured jitter buffer
109 latency value; iii) adjusts the jitter buffer latency value to a minimum preconfigured jitter
110 buffer latency value if the target delay value is less than the minimum preconfigured
111 jitter buffer latency value; iv) decrements the jitter buffer latency value by a
112 predetermined maximum decrement value if adjusting the jitter buffer latency value to

10/826,204

Marked to Show Changes

113 the target delay would result in decrementing the jitter buffer latency value by more than
114 the predetermined maximum decrement value; and v) increments decrements the jitter
115 buffer latency value by a predetermined maximum increment value if adjusting the jitter
116 buffer latency value to the target delay would result in incrementing the jitter buffer
117 latency value by more than the predetermined maximum increment value.

118 The histogram module may: i) calculate a histogram value from each delay value;
119 ii) store each histogram value in a bin or a sub-gram associated with the ~~then~~-current
120 jitter buffer latency value; and iii) calculate the target delay value upon completion of the
121 sub-gram. The sub-gram may be a logical portion of a histogram memory comprising a
122 predetermined quantity of logical bins. The sub-gram is considered complete when the
123 predetermined quantity of histogram values have been stored in the sub-gram (e.g. the
124 bins are full).

125 The histogram module may calculate the target delay value by: i) determining a
126 low value which the predetermined portion of the histogram values are less than the low
127 value and the remainder of the histogram values are greater than the low value; and ii)
128 setting the target delay value to the difference between zero and the low value.

129 In addition, the histogram module may calculate a quantity of frames that must
130 be added or dropped to compensate for a discontinuity in the output time stamp
131 sequence caused by the adjustment in the jitter buffer latency value. The histogram
132 module may add the value of the jitter buffer latency to the low value to generate a
133 resulting value. If the resulting value is greater than zero, a quantity of frames equal to
134 the resulting value divided by the output time stamp increment are dropped from the
135 jitter buffer. If the resulting value is less than zero, a quantity of frames equal to (the
136 absolute value of) the resulting value divided by the output time stamp increment are
137 created and added to the jitter buffer.

138 For a better understanding of the present invention, together with other and
139 further aspects thereof, reference is made to the following description, taken in

10/826,204

Marked to Show Changes

140 conjunction with the accompanying drawings. The scope of the present invention is set
141 forth in the appended claims.

142

143

144

Brief Description of the Drawings

145 Figure 1 is a block diagram representing a system for providing VoIP
146 communication services over a frame switched network in accordance with one
147 embodiment of the present invention;

148 Figure 2 is a block diagram representing an exemplary VoIP frame in accordance
149 with one embodiment of the present invention;

150 Figure 3 is a flow chart representing exemplary operation of an output time stamp
151 index module in accordance with one embodiment of the present invention;

152 Figure 4 is a flow chart representing exemplary operation of a delay calculation
153 circuit in accordance with one embodiment of the present invention;

154 Figure 5 is a flow chart representing exemplary operation of a histogram module
155 in accordance with one embodiment of the present invention; and

156 Figure 6 is a table representing exemplary configuration values for the jitter buffer
157 system of the present invention.

158

159

Detailed Description of the Exemplary Embodiments

160 The present invention will now be described in detail with reference to the
161 drawings. In the drawings, each element with a reference number is similar to other
162 elements with the same reference number independent of any letter designation
163 following the reference number. In the text, a reference number with a specific letter
164 designation following the reference number refers to the specific element with the
165 number and letter designation and a reference number without a specific letter
166 designation refers to all elements with the same reference number independent of any
167 letter designation following the reference number in the drawings.

10/826,204

Marked to Show Changes

168 It should also be appreciated that many of the elements discussed in this
169 specification may be implemented in a hardware circuit(s), a processor executing
170 software code, or a combination of a hardware circuit(s) and a processor or control
171 block of an integrated circuit executing machine readable code. As such, the term
172 circuit, module, server, or other equivalent description of an element as used throughout
173 this specification is intended to encompass a hardware circuit (whether discrete
174 elements or an integrated circuit block), a processor or control block executing code, or
175 a combination of a hardware circuit(s) and a processor and/or control block executing
176 code.

177 Figure 1 represents a voice over Internet Protocol (VoIP) system 10 which
178 includes a terminal adapter 14 useful for implementing the improved adaptive jitter
179 buffer system 36 of the present invention. Although the improved adaptive jitter buffer
180 system 36 is implemented within a terminal adapter 14 for purposes of illustrating the
181 invention, it should be appreciated that the invention is useful in conjunction with other
182 packet audio systems such as voice over Internet Protocol (VoIP) telephones.

183 The system 10 comprises a frame switched network, such as the Internet 12,
184 interconnecting a plurality of VoIP telephony endpoints such as the terminal adapter 14
185 and a remote endpoint 46. In the exemplary embodiment, the terminal adapter 14 is
186 coupled to a traditional PSTN telephone device 16 and a local area network 52. The
187 local area network 52 in turn is coupled to an ISP network 48 (which is a part of the
188 Internet 12) by an ISP gateway 50. In exemplary embodiments, the ISP network 48 and
189 gateway 50 may be: i) a hybrid fiber/cable network and cable modem respectively; ii) a
190 telephony service provider network and digital subscriber line (DSL) modem
191 respectively, or iii) other known networking technologies for providing Internet services
192 to a customer's premises.

193 In operation, the terminal adapter 14 emulates a central office switch at a PSTN
194 port 34 for providing telephone service to the PSTN device 16 coupled thereto. The
195 PSTN telephone service may be provided utilizing traditional PSTN analog or digital call

10/826,204

Marked to Show Changes

196 signaling and voice band communications. The terminal adapter 14 further links the
197 PSTN call signaling and voice band communications (e.g. a PSTN call leg) with VoIP
198 call signaling and media session communications (e.g. a VoIP call leg) over the Internet
199 12 to the remote endpoint 46 to facilitate a telephone call between the PSTN device 16
200 and the remote endpoint 46.

201 Turning briefly to Figure 2, an exemplary media session frame 60 for transporting
202 compressed digital audio over the Internet 12 is shown in block diagram form. The
203 media session frame 60 comprises an IP header 62, a UDP header 64, an RTP header
204 66, and audio samples 72 which are compressed digital audio data representing a
205 discrete portion of the voice band.

206 The IP header 62 comprises such information as the source IP address from
207 which the frame 60 was generated and destination IP address to which the frame 60 is
208 to be routed over the Internet 12. The UDP header 64 comprises such information as
209 the port number which identifies the source and destination applications. The RTP
210 header 66 comprises such information as a sequence number 68 and a transmission
211 time stamp 58. The sequence number 68 defines the frame's sequence or position
212 amongst a plurality of other frames generated by the framing module 24. The
213 transmission time stamp 58 represents a time at which the frame was generated. The
214 difference between transmission time stamp values of sequential frames represents the
215 period between frames and also should approximate the duration of time represented
216 by the audio samples 72 within the frame. The transmission time stamp 58 and the
217 sequence number 68 provide information needed for re-generation of voice band at the
218 receiving VoIP device even though transmission latency time for each frame 60 may
219 vary randomly from that of other frames 60.

220 Returning to Figure 1, the terminal adapter 14 may comprise a VoIP client 20, a
221 dialog system 22, jitter buffer system 36, as well as each of a network interface 18 for
222 coupling to the local area network 52 and a PSTN FXO port 34 for coupling to the
223 traditional PSTN telephony device 16.

10/826,204

Marked to Show Changes

224 The network interface 18 utilizes known physical layer systems which are
225 compliant with those utilized by the local area network 52 and known internet protocol
226 systems (typically referred to as an "IP Stack") for communicating with remote IP
227 endpoints over the local area network 52. In the exemplary embodiment, the physical
228 layer systems of the network interface 18 may operate a known communication
229 standard such as USB or Ethernet for communicating with the ISP gateway 50.

230 In operation, the network interface 18 receives session set up frames from the
231 VoIP client 20 and media session frames from the dialog system 22, packages the
232 frames as UDP/IP frames with applicable source and destination socket information,
233 and forwards the UDP/IP frames to the applicable remote device over the local area
234 network 52. The network interface 18 also receives UDP/IP frames over the local area
235 network 52 and presents the data therein to either the VoIP client 20 or the dialog
236 system 22 based on a destination socket (IP address and port number) of the received
237 frame.

238 The VoIP client 20 may operate known VoIP signaling systems such as: i) the
239 Media Gateway Control Protocol (MGCP, RFC3435, RFC3661) for exchanging call set
240 up messages with a call agent (not shown), gateway (not shown) and/or the remote
241 endpoint 46; or ii) the Session Initiation Protocol (SIP) for exchanging call set up
242 messages with a SIP compliant proxy server (not shown) and/or the remote endpoint
243 46.

244 The VoIP client 20 also includes circuits for exchanging call signaling and
245 session status signals 52 with the dialog system 22 such that the dialog system 22 may
246 exchange corresponding call signaling and session status signals (such as ringing,
247 busy, and MGCP caller ID messages) with the PSTN device 16 as known analog or
248 digital signaling appropriately modulated onto the PSTN link 53.

249 The dialog system 22 may be embodied in a digital signal processing (DSP)
250 circuit and may include a PSTN driver module 32, a signaling module 30, a
251 decompression module 28, a compression module 26, and a framing module 24.

10/826,204

Marked to Show Changes

252 The PSTN driver 32 is coupled to each of the signaling module 30, the
253 compression module 26, the decompression module 28, and the PSTN device 16 (via a
254 PSTN port 34 to which the PSTN device 16 is coupled). In operation the PSTN driver
255 32 emulates a central office switch for providing telephone service to the PSTN device
256 16 over the link 53. More specifically, the PSTN driver 32 detects a voice band signal
257 generated by the PSTN device 16 (e.g. local voice band), samples the signal at 800Khz
258 to generate a digital audio signal, and provides the digital audio signal to each of the
259 signaling module 30 and the compression module 26. With respect to voice band
260 generated by the remote endpoint 46 (remote voice band), the PSTN driver receives a
261 digital audio signal from the decompression module 28 (representing the remote voice
262 band) and recreates PSTN analog or digital voice band for coupled to the PSTN device
263 16.

264 The signaling module 30: i) receives the digital representation of the local voice
265 band from the PSTN driver 32; ii) utilizes pattern matching techniques to detect
266 traditional tone call signaling within the local voice band such as DTMF tones, and
267 provides corresponding signals 54 to the VoIP client 20 such that the VoIP client 20 can
268 generate corresponding VoIP messages for transmission to the applicable endpoint
269 over the network 12.

270 The signaling module 30 further receives signals 54 from the VoIP client 20
271 (corresponding to VoIP messages received by the VoIP client 20) and provides
272 corresponding digital signaling to the PSTN driver 32 such that the PSTN driver can
273 appropriately modulate corresponding voice band signaling (e.g. in-band signaling) such
274 as dial tone, DTMF tones, ring back signal, busy signals, call waiting signal, caller ID
275 signals, and flash signals on the PSTN link 53.

276 The compression module 26 receives the digital representation of the local voice
277 band from the PSTN driver 32 and operates algorithms which compress the local voice
278 band into compressed digital audio samples 78. The audio samples 78 are linked to the
279 framing module 24 which: i) packages such samples into a real time protocol (RTP)

10/826,204

Marked to Show Changes

280 stream of frames; and ii) presents each frame of the RTP stream to the network
281 interface circuit 18 for packaging as a UDP/IP frame 60 for transport to the destination
282 endpoint.

283 The decompression module 28 receives from the jitter buffer system 36, each
284 frame of an RTP stream of compressed digital audio generated by the remote endpoint
285 46, in response to generating a clock signal 61 therefore. The decompression module
286 28 further decompress the audio samples within each frame 60 to re-generate a digital
287 representation of the remote voice band – which in turn is coupled to the PSTN driver
288 32 for recreation of the remote voice band on the PSNT link 53.

289 Exemplary compression/decompression algorithms utilized by the compression
290 module 26 and the decompression module 28 include: i) algorithms that provide minimal
291 (or no) compression (useful for fax transmission) such as algorithms commonly referred
292 to as G.711, G.726; ii) very high compression algorithms such as algorithms commonly
293 referred to as G.723.1 and G.729D; and iii) algorithms that provide compression and
294 high audio quality such as algorithms commonly referred to as G.728, and G. 729E.

295

296 **Jitter Buffer System**

297 As discussed in the background, a problem with VoIP telephony is that, even
298 though frames may be transmitted in sequence and at regular periods, the frames may
299 arrive at the destination endpoint both out of sequence and with variations in their
300 transport times (e.g. the time it takes for each frame to be routed from its source to its
301 destination over the Internet will vary).

302 To enable the voice band of the remote endpoint 46 to be regenerate at the
303 terminal adapter 14, a jitter buffer system 36 is used for buffering frames. In operation,
304 the jitter buffer system 36 corrects for variations in transport time between frames.
305 More specifically, the jitter buffer system 36 receives each frame sent by the remote
306 endpoint 46, stores each received frame in a jitter buffer 44, and sequentially releases
307 frames from the jitter buffer 44 to the decompression module 28 of the dialog system 22

10/826,204

Marked to Show Changes

308 at a release period time corresponding to a clock signal 61 provided by the
309 decompression module 28 (discussed herein).

310 Effectively, the jitter buffer 44 generates an additional delay (e.g. buffer delay)
311 between receiving the frame and release of the frame to the dialog system 22 such that
312 the Internet transport delay plus buffer delay (collectively jitter buffer latency) is
313 generally a fixed latency for all frames.

314 It should be appreciated that if the jitter buffer latency is a small value, certain
315 frames with a high transport delay will have a transport delay greater than the jitter
316 buffer latency and will be unusable. These are lost frames which are dropped and
317 result in degradation of quality of the re-created voice band.

318 It should also be appreciated that if the jitter buffer latency value is large,
319 although frames may not be dropped, the large jitter buffer latency may create a delay
320 of the re-created voice band noticeable to the user.

321 The jitter buffer system 36 comprises a jitter buffer 44, a output time stamp index
322 module 42, a delay calculation module 38, and a histogram module 40. In operation,
323 the output time stamp index module 42 calculates an output time stamp 59 after
324 receiving each jbLatency value 55 (which are periodically provided by the histogram
325 module 40). The output time stamp 59 will be a transmission time stamp 58 of the first
326 frame received following the reset plus the jbLatency value 55. Thereafter, the output
327 time stamp 59 is incremented by a fixed value each time a frame is released to the
328 decompression module 38 in response to the clock signal 61. The fixed value may be
329 referred to as the increment.

330 The delay calculation circuit 38 calculates a delay value 63 for each received
331 frame 60 by subtracting the transmission time stamp 58 from the output time stamp
332 value 59. If the result is either negative (e.g. under-run) or greater than a
333 predetermined maximum allowable delay, the delay calculation circuit 38 drops the
334 frame. (e.g. either the frame is not written to the jitter buffer 44 or is removed from the
335 jitter buffer 44).

10/826,204

Marked to Show Changes

336 In addition, if the delay calculation circuit 38 detects a significant change in the
337 value of transmission time stamp 58 between sequential frames, the delay calculation
338 circuit 38 will generate a reset signal 57 to the histogram module 40 to force the
339 histogram module 40 to provide a new value of jbLatency 55.

340 The histogram module 40 provides an initial value of jbLatency 55 to the output
341 time stamp index module 42 upon start of a sequence for frames and upon the output
342 time stamp index module 42 providing a reset signal 57.

343 Further, the histogram module 40 receives each delay value 63 and calculates a
344 histogram value for storage in a histogram memory 41. The histogram value is equal to
345 the delay value 63 minus the ~~then~~ current value of jbLatency 55 – or, stated another
346 way, the histogram value is a normalized delay value that would have been the delay
347 had the value of jbLatency been zero.

348 The histogram memory 41 may be a storage system utilized to represent a
349 plurality of graphical histograms – each referred to as a sub-histogram. The sub-
350 histogram includes a fixed quantity of sequential histogram values which correspond to
351 a single value of jbLatency 55. When a sub-histogram reaches its limit of values, that
352 sub-histogram is considered complete and a new sub-histogram is started.

353 Following completion of each sub-histogram a new value of jbLatency 55 may be
354 calculated based on the histogram values stored in the most recent predetermined
355 number of sub-histogram completed. The new value of jbLatency 55 is provided to the
356 output time stamp index module 42 such that it may again calculate an initial value of
357 output time stamp 59 (e.g. when reset). In addition, frames within the jitter buffer 44
358 may be added or dropped to compensate for the adjustment of jbLatency 55 – or, stated
359 another way, frames may be created or dropped to accommodate a the adjustment in
360 the buffer delay. It should be appreciated that the created frames can not be real audio
361 data, but comprise filler audio data or audio data extrapolated from adjacent (in time)
362 frames to provide hardware of the decompression module 28 with compressed audio
363 data on a periodic basis while still adjusting time.

10/826,204

Marked to Show Changes

364 Except for threshold limitations and minimum/maximum value limitations, the new
365 value of *jbLatency* 55 is a delay value which, if it had been used as the *jbLatency* value
366 55 during the histogram period, would have resulted in a predetermined portion of the
367 frames being dropped. More specifically, a configuration value known as *dropsPerMil*
368 218 (Figure 6) may be the predetermined portion of frames expressed in a quantity of
369 frames per one-thousand frames.

370 The flow chart of Figure 3 represents exemplary operation of the output time
371 stamp index 42. The two inputs of the output time stamp index 42 are a *jbLatency* value
372 55 from the histogram module 40 and a clock signal 61 from the decompression module
373 28. Step 140 represents an event loop waiting for on of those two inputs.

374 In the event that a *jbLatency* value 55 is received, the output time stamp index 42
375 calculates an initial value of output time stamp 59 by setting output time stamp 59 equal
376 to the value of *jbLatency* 55 plus the value of transmit time stamp 58 of the next
377 received frame (or the most recently received frame). Calculation of output time stamp
378 59 is represented by box 142 and after performing the calculation, the output time
379 stamp index 42 returns to the event loop 140.

380 In the event that a clock signal 61 is received, the output time stamp index 42
381 increments the value of output time stamp 59 by the fixed increment at step 144 – and
382 thereafter returns to the event loop 140.

383 The flow chart of figure 4 represents exemplary operation of the delay calculation
384 module 38. Step 146 represents receiving a frame 60 from the network interface circuit
385 18.

386 Step 148 represents determining whether there has been a significant change in
387 the value of transmission time stamp 58. If the change in the value of transmission time
388 stamp 58 between frames is significant it should be appreciated that a significant
389 discontinuity exists. As such, a reset signal 57 is generated at step 150 such that a new
390 value of *jbLatency* 55 can be calculated and a new value of output time stamp 59 can
391 be calculated.

10/826,204

Marked to Show Changes

392 Alternatively, if there is not a significant change in the value of transmission time
393 stamp 58, the delay calculation circuit sets the value of delay 63 equal to the value of
394 transmission time stamp 58 of the frame 60 less the value of the output time stamp 59
395 at step 154.

396 If delay 63 is less than zero, as represented by decision box 156, or greater than
397 a preconfigured value known as maxDelay 210 (Figure 6) as represented by decision
398 box 160, the frame 60 is dropped as represented by boxes 158 and 162 respectively.

399 If the frame 60 is not dropped, the frame (and its transmission time stamp value
400 58) are written to the jitter buffer 44 at step 164 and the value of delay 63 is provided to
401 the histogram module 40.

402 The flow chart of Figure 5 represents exemplary operation of the histogram
403 module 40 which, as discussed, periodically generates the values of jbLatency 55 used
404 by the output time stamp index module 42 to assign values of output time stamp 59 to
405 received frames.

406 Step 102 represents setting the value of jbLatency 55 equal to the value of
407 initialLatency 208. The value of initialLatency 208 is a configurable parameter stored in
408 the configuration value table 200 (Figure 6). Following step 102, the histogram module
409 40 enters a loop defined by steps 104 through 134 in which it periodically updates the
410 value of jbLatency 55 until such time as a reset signal 57 is received.

411 Step 104 represents obtaining a value of delay 63 from the delay calculation
412 circuit 38 and step 106 represents generating a histogram value 136. As discussed, the
413 histogram value 136 is equal to the value of delay 63 less the value of jbLatency 55.
414 Step 108 represents storing the histogram value 136 in a bin of a sub-gram of the
415 memory 41.

416 As previously discussed, the value of jbLatency 55 is periodically updated using
417 a histogram of delay values (e.g. output time stamp 59 less transmission time stamp 58)
418 to select a target delay that would have resulted in a predetermined portion of the
419 frames being dropped.

10/826,204

Marked to Show Changes

420 The problem is that the value of output time stamp 59 is itself effected by the
421 value of jbLatency 55 that was in use at the time the frame was written to the jitter buffer
422 44. Therefore, if delay values were used for a histogram, an iterative approach to
423 determine target delay would be required. More specifically, the system would have to
424 select a trial target delay, adjust all of the delay values to what the delay value would
425 have been if jbLatency had been set to the trial target delay, determining the portion of
426 frames that would have been dropped, and then re-adjusting the trail target delay.

427 Therefore, rather than storing the value of delay 63 in the histogram memory 41,
428 the histogram value 136 is stored in a bin of a sub-gram. The sub-gram comprises a
429 predetermined plurality of bins defined by the value of bin 216 in the configuration value
430 table 200 (Figure 6). Because the value of jbLatency 55 is only adjusted upon
431 completion of a sub-gram, all histogram values 136 in the sub-gram are calculated
432 using the same value of jbLatency 55. Further, to determine a target delay, all of the
433 delay values have already been normalized by subtracting the ~~then~~ current value of
434 jbLatency. Or stated another way, the histogram value is what the delay value would
435 have been had the jbLatency value been zero at the time the frame was written to the
436 jitter buffer 44.

437 Step 110 represents determining whether the sub-gram is complete. If the sub-
438 gram has stored a predetermined number of values, it is complete. If not, the histogram
439 module 40 returns to step 104 to again receive a value of delay 63 from the delay
440 calculation module 38. When the sub-gram is complete, the delay calculation
441 module 40 performs steps 112 through 134 for updating the value of jbLatency 55. Step
442 112 represents determining a target delay value. As discussed, the target delay value
443 is the value which, if used as the value of jbLatency 55 for the most recently received
444 predetermined quantity of frames (e.g. the frames of the most recently predetermined
445 number of sub-grams – with the predetermined number of sub-grams being equal to the
446 value of grams 214 in the configuration value table 200) would have resulted in a
447 predetermined portion of the frames being dropped. The predetermined portion of

10/826,204

Marked to Show Changes

448 frames (in units of frames per thousand) being equal to the value of dropsPerMil 218 in
449 the configuration table.

450 More specifically, the target delay value is determined by selecting the histogram
451 value that results in the predetermined portion of the histogram values 136 being less
452 than the histogram value. This histogram value can be called the low value. The target
453 delay value is then the difference between zero and this low value. It should be
454 appreciated that because the value of dropsPerMil 218 will be a small percentage of all
455 frames and because the histogram values have been normalized to values (as if
456 jbLatency had been zero), the low value will always be less than zero.

457 To determine the adjustment to jbLatency 55, a target value is calculated. The
458 target value is the absolute value of the low value – or stated another way, the
459 difference between zero and the low value.

460 In addition, an adjustment in the value of jbLatency 55 will cause a discontinuity
461 in the value of output time stamp 59 between sequential frames. Frames must be
462 deleted or added to accommodate this discontinuity. To determine deletion or addition
463 to frames, the following calculations are used. First, the value of jbLatency 55 is added
464 to the low value. If the resulting value is greater than zero, frames equal to the resulting
465 value divided by the output time stamp increment must be dropped. Similarly, if the
466 resulting value is less than zero, frame equal to the resulting value divided by the output
467 time stamp increment must be created and added to the jitter buffer 44.

468 In general, the value of jbLatency 55 will be updated to the target delay value
469 (and the appropriate frame adjustments made) at step 130. However, in certain
470 instances, there are practical limits on adjustments that should be made to the value of
471 jbLatency 55. Steps 114, 118, 122, and 126 represent testing the practical limits. More
472 specifically, at step 114, if the change in the value of jbLatency 55 would be less than a
473 predetermined hysteresis value (e.g. value of hysteresis 220 in the configuration value
474 table 200), the value of jbLatency 55 is not changed as the change would be too small
475 and result in too frequent of adjustments that are not really necessary.

10/826,204

Marked to Show Changes

At step 116, if the target delay is greater than a predetermined maximum delay (e.g. value of maxDelay 210 in the configuration value table 200), the value of jbLatency 55 is set to maxDelay 210 (and the frames adjusted accordingly) at step 117 rather than the target delay. Typically maxDelay 210 will be selected as a value that is on the threshold of noticeable delay to the user and greater frame loss would be more tolerable to the user than greater delay.

At step 118, if the target delay is less than a predetermined minimum jbLatency (e.g. value of minLatency 208 in the configuration value table 200), the value of jbLatency 55 is set to minLatency 208 (and the frames adjusted accordingly) at step 119 rather than the target delay.

At step 120, if the adjustment in jbLatency 55 will be a decrement greater than a predetermined maximum decrement (e.g. value of maxDrops 222 in the configuration value table 200), then the value of jbLatency 55 will be decremented by maxDrops 222 at step 121 rather than to the target delay.

At step 120, if the adjustment in jbLatency 55 will be a decrement greater than a predetermined maximum decrement (e.g. value of maxDrops 222 in the configuration value table 200), then the value of jbLatency 55 will be decremented by maxDrops 222 rather at step 121 rather than to the target delay.

At step 122, if the adjustment in jbLatency 55 will be an increment greater than a predetermined maximum increment (e.g. value of maxAdds 224 in the configuration value table 200), then the value of jbLatency 55 will be incremented by maxAdds 222 at step 123 rather than to the target delay.

If none of the limits are reached at tests 114-122, then at step 130, the value of jbLatency 55 is updated to the target delay. Step 132 represents starting a new sub-gram into which histogram values will be stored and dropping the oldest sub-gram from future calculations of target latency. The configuration value table 200 includes a value of grams 214 which defines the quantity of sub-grams used for calculating jbLatency 55. Because a new sub-gram has been created, the oldest is dropped to assure that the

10/826,204

Marked to Show Changes

504 quantity of sub-grams used for calculation remains at the value of grams 214.
505 Step 134 comprises providing the value of jbLatency 55 to the output time stamp
506 index 42 and associating the value of jbLatency 55 with the new sub-gram.

507 It should be appreciated that the systems and methods discussed herein provide
508 for a jitter buffer system which corrects for variations in transport time between frames –
509 and more specifically dynamically adjusts jitter buffer latency based on histogram
510 characteristics to target a frame loss value that optimizes audio degradation trade off
511 between excessive frame loss and excessive latency.

512 Although the invention has been shown and described with respect to certain
513 preferred embodiments, it is obvious that equivalents and modifications will occur to
514 others skilled in the art upon the reading and understanding of the specification. The
515 present invention includes all such equivalents and modifications, and is limited only by
516 the scope of the following claims.

517

518

10/826,204

Marked to Show Changes

Abstract of the Disclosure

A jitter buffer system for reducing jitter in a packet audio reception device comprises an output time stamp index, a delay calculation module, and a histogram module. The output time stamp index determines an initial output time stamp value upon receipt of a jitter buffer latency value and increments the output time stamp upon release of each of the frames by a jitter buffer. The delay calculation module calculates a delay value for each of the sequence of received frames. ~~The delay value is equal to the time difference between the output time stamp value and a transmission time stamp assigned to the frame by the transmitting system.~~ The histogram module is coupled to each of the output time stamp index and the delay calculation module, the histogram module periodically calculates a target delay value which, based on a buffered history of values representing the delay value of each of a fixed quantity of the most recently received frames ~~from the sequence of receive frames~~, would have resulted in a predetermined portion of the fixed quantity of frames being dropped, and adjusting the jitter buffer latency value to a value equal to the target value.